Research IPOR Report

TWO NEW APPROACHES TO FACTOR ANALYSIS

ANNUAL TECHNICAL REPORT

JUNE, 1953

INTERNATIONAL PUBLIC OPINION RESEARCH, INC.
EMPIRE STATE BUILDING, NEW YORK CITY

TWO NEW APPROACHES TO FACTOR ANALYSIS

Prepared for:

Office of Naval Research

By

International Public Opinion Research, Inc. &

Israel Institute of Applied Social Research

June, 1953

ANNUAL TECHNICAL REPORT

Contract No. Nonr - 731(00)

TWO NEW APPROACHES TO FACTOR ANALYSIS

This project has been concerned during the past year with elaborating and empirically testing some new approaches to the problem of factor analysis. Three kinds of content have been studied:

(a) mental ability tests

(b) personality schedules and ratings

(c) attitude questionnaires

With respect to contents of types (a) and (b) above, tables of product-moment correlations were culled from those published in the scientific literature. Many of these had already been subjected to older types of factor analyses (usually in the sense of Thurstone). The work of the project was to analyze (or re-analyze) these data according to the new approach called that of the <u>radex</u>.

With respect to the attitudinal data (c) above, all the available tabulations from all the researches of the Israel Institute of Applied Social Research were scrutinized, and those that seemed most appropriate on certain <u>a priori</u> grounds (to be described below) were selected for an intensive new type of analysis. These tabulations were unpublished, and indeed are of the type that are rarely published — even in the most extended reports -- by research organizations, since they are so bulky and deal with what are usually considered too minute details.

The basic attitudinal tabulations needed are the complete scattergram of one scale against another, with around ten distinct ranks (scores or intervals) for each scale separately. A further requirement is that the total sample be large enough so that each rank have an average of not much less than one hundred cases. A further and crucial desideratum is that each scale separately be unidimensional in the scalogram sense, so that it have psychologically meaningful principal components. It is the polytone and multivalued relationship among attitudes -- as theoretically generated when they have a principal component in common — that was studied empirically in this part of the project.

A major technical difference here between the new factor theories, then, is that the radex was studied only with respect to variables whose mutual regressions were supposedly linear (hence univalued and monotone), and where in particular product-moment correlations were sufficient parameters.

This type of relationship seems to hold among scores on current mental tests. In contrast, a fundamental point of departure our project makes with respect to attitude data is that it rejects the hypothesis of linearity --- or more generally of monotonicity and single-valuedness. In effect, our project's findings support clinical psychologists' objections to the use of conventional statistical techniques in trying to study the structure of attitudes, or to find out what attitudes have in common. Our empirical results verify that in many cases polytone and multivalued regression systems hold between two different scalable attitudes. This has been predicted by the scalogram theory, and is now amply verified empirically (for the first time) in the present project. Complete scattergrams are needed for this purpose; correlation coefficients and the like are often worse than useless (they are usually close to zero), for they distract attention from the more fundamental aspects of multivaluedness.

The beginnings of the theory of the radex was developed by the project leader, Louis Guttman, some four years ago, in the form of the special cases of the simplex and the circumplex. This theory was presented orally by him in a lecture tour of leading American universities in the spring of 1951. Only two or three fairly good empirical examples of simplexes were available then, and only one (somewhat dubious) empirical example of a circumplex.

The present ONR project began actual work in January, 1952, and has steadily discovered example after example of well-tested simplexes and circumplexes. The term radex was invented in the middle of the project in order to help concretize the full outline that was emerging from the work.

Halfway through the project, the Director had occasion to write a rather detailed exposition of both the theory and empirical verification of radex theory. This was nominally to be only a written version of his previous oral lectures in the United States. In fact, it is an almost completely new, and greatly expanded, presentation, based on the empirical materials already completed to that time by the present project. This monograph of over 100 manuscript pages will appear in the fall of 1953 in a volume edited by Paul F. Lazarsfeld, entitled Mathematical Thinking in the Social Sciences, and will constitute the first scientific publication resulting from this project. Some extensions of the theory presented there are already necessary, as will be indicated in the present report.

The multivalued relationships among attitudes require an entirely new departure in computing techniques. No existing methods of determining regressions are appropriate, since they all assume single-valuedness in at least one direction. Over half of the project, in this direction, was spent in trying to perfect the theory and practice of computing. After laborious explorations into numerous blind alleys, two standard and complementary computing procedures emerged. Neither by itself is

sufficient for all occasions. It seems that it can be argued on theoretical (purely mathematical) grounds that no one computing system will succeed under any and all circumstances. The pair of procedures we have developed are quite workable in practice, and seem quite safe, especially when one is used to check the other. One is called the technique of quantiles, and the other of latent vectors. They tackle the problem in an entirely different manner from each other.

The second of the computing procedures requires latent vectors (and roots) of certain matrices. Our matrices are of larger order than those usually handled by physicists and others who do similar computing work in practice on matrices. Existing numerical procedures even for this standard problem of physicists and mathematicians proved too laborious for our data. An excessive amount of labor and time was needed to get even the simplest numerical answer for only one of our cross-tabulated attitudes. The Project Director was able finally to develop new procedures for computing latent roots and vectors which seem far more rapid than any as yet available. These will be published in the Annals of Mathematical Statistics in a paper by Louis Guttman entitled: "New Direct and Iterative Techniques for Computing Latent Vectors and/or Roots". This will be the second scientific publication resulting from this project. While this particular paper has no direct sociological or psychological content, its results were essential for effective furtherance of an important development of the attitudinal part of the project.

The ONR may find it useful to call attention of this paper on latent vectors and roots not only to those associated with the Human Relations Branch, but also to other branches -- especially in physics, chemistry, applied mathematics, and computing -- where the same type of numerical problem arises and where the new computing procedures of our paper may prove helpful.

In the next part of this report, we shall summarize the year's findings with respect to the radex approach to factor analysis. The following part will similarly summarize the year's findings with respect to inter-attitude relationships. The three Appendices state, respectively, the original purpose of the project, the personnel employed, and the breakdown of costs for the entire year.

Much of the data analyzed in this project lend themselves already to be written up in formal fashion for scientific publication. A series of articles is being planned for submission to the scientific literature. The problem is also being weighed whether to write a full-length monograph on the radex, to complement that already in press. This will depend in part on the extension of the project to a second year. The first year has made quite clear the fact that further fundamental progress in this direction of factor analysis requires assembling original data designed directly to test the radex theory. Data published in the literature have now been used to the extent they can be for this purpose, but they are not fully adequate for the complete picture, giving only parts a bit at a time.

In an extension to a second year, it would be highly desirable to reassemble a comprehensive battery of mental tests, according to the streamlining made possible by radex theory, and administer it to large samples of different age groups. This would enable a full and direct test of radex theory as a whole, instead of simplexes and circumplexes separately as in the past year's work, and simultaneously provide the parsimoneous batteries for practical prediction use according to the powerful predictive properties of the radex structure.

Results of the Project with Respect to the Radex

1. Perhaps the most important finding of the project is that it provides ample empirical evidence, on the basis of materials assembled by other investigators with other purposes in mind, that the radex approach to factor analysis is appropriate for many fields of mental testing, and some aspects of personality assessment. It gives a picture of the *functional* nature of the interrelationships among different abilities and traits, in a manner not made possible by previous approaches.

The conventional multiple common factor approaches of Thurstone, Burt, and others usually posit that a relatively small number of common-factors can explain the observed intercorrelations of test data. From the statistical techniques used to estimate these hypothesized common-factors, it is clear that they are (in the limit) but arithmetic averages of the observed tests (cf. Louis Guttman, "Multiple Group Methods for Common-Factor Analysis: Their Basis, Computation, and Interpretation," *Psychometrika*, June, 1952,17:209-222). In these theories, just as the tests are hypothesized to be weighted sums of the factors, just so are the factors essentially weighted sums of the tests. Therefore, in a sense, the common-factors of these theories are on the same level as the tests.

It is not customary in science to "explain" phenomena by means of phenomena on the same level, but an explanation on a lower level is sought. Thus, "water" is explained by "hydrogen" and "oxygen", but neither of the latter elements is a function in turn of the compound "water".

In radex theory, it is *not* hypothesized that a small number of common-factors exists that can be determined as averages of tests. Instead, it assumes indefinitely many (for convenience, even continuation of) common-factors to operate in a test battery, which under certain circumstances give rise to orderly arrangements among test inter-correlations. These orderly arrangements are observable without any specific theory as to the nature of the common-factors, and without any calculations aimed at "locating" the common-factors.

The general theory of the radex has been described in the monograph resulting largely from this project (Louis Guttman, "A New Approach to Factor Analysis: The Radex," in Paul F. Lazarsfeld, Mathematical Thinking in the Social Sciences, in press), with special emphasis on two particularly simple order systems: the simplex and the circumplex.

One of the surprises of the project's results is the relatively large number of approximate simplexes and circumplexes found in the existing scientific literature. The simple order patterns to be found in their data have been overlooked apparently by all previous investigators.

2. A corollary of the finding that many empirical data do conform to radex theory is that future testing programs can be made more efficient for prediction purposes and at the same time more economical. Smaller batteries of tests can be designed according to the observed order patterns, and yield better predictions than existing non-ordered batteries. To achieve this goal, of course, requires research outside the scope of the first year's work of this project, but the work completed to date is empirical evidence that the goal is very approachable.

3. A brief description of radex theory is given by the following quotation from the monograph in press:

> "A set of variables whose intercorrelations conform to the general order pattern prescribed by the new theory will be called a radex. This is a word designed to indicate a 'radical expansion of complexity'.
>
> "Two distinct notions are involved in a radex. One is that of a difference in kind between tests, and the other is that of a difference in degree. Each of these notions will give rise to a separate concept of order among tests, so that a radex is ultimately at least a doubly-ordered system. In this monograph, we shall treat only the simplest case of the radex which can be completely portrayed by a simple, two-dimensional diagram. The empirical evidence thus far -- as I shall exhibit shortly -- shows that mental test data can be surprisingly well accounted for by such a diagram.
>
> "Within all tests of the same kind, say of numerical ability, differences will be in degree. We shall see that addition, subtraction, multiplication, and division differ among themselves largely in the degree of their complexity. Such a set of variables will be called a simplex. It possesses a simple order of complexity. The tests can be arranged in a simple rank order from least complex to most complex.

"Correspondingly, all tests of the same degree of complexity will differ among themselves only in the kind of ability they define. We shall postulate a law of order here too, but one which is not from 'least to 'most' in any sense. It is an order which has no beginning and no end, namely, a circular order. A set of variables obeying such a law will be called a circumplex, to designate a 'circular order of complexity'. Our empirical data will testify that different abilities such as verbal, numerical, reasoning, etc. do tend to have such an order among themselves.

"In the more general case, tests can differ among themselves simultaneously both in degree and in kind of complexity, and the general structure here is the radex. Thus, within a radex, one can usually isolate simplexes by keeping the content of the abilities constant and by varying the degree of complexity; and one can also usually isolate circumplexes by keeping degree of complexity constant and then varying the content.

"In practice, in this present first effort, we must begin without knowledge of the ultimate map of the radex. So we begin by studying separately defined universes of content and see if each is an approximate simplex. Then by selecting one test from each universe, we see if a circumplex emerges. This monograph is devoted to giving examples of this procedure. On a later occasion, when some current empirical work on the radex is completed, I expect to present a fairly detailed radex map of certain human abilities.

"A powerful feature of a simplex structure is its immediate use for prediction purposes. The same holds for a circumplex structure, as well as for the more general radex. This feature seems largely lacking in previous theories of factor analysis. Radex theory opens a clear path to better predictions with less tests."

4. The practical procedure used in this project to test a given set of data to see if they form a simplex is illustrated by the following example. This example is also a highly practical one for showing how to get better predictions — and with less work — than are being obtained at present from an internationally popular current test battery.

Radex theory has recently interested the Psychological Testing Unit of the Israel Defence Army. At the December, 1952, annual national meetings of the Israel Psychological Association, Mr. K. Reuben Gabriel reported the following results from the Unit's data. Raven's Progressive Matrices

Test, consisting of five subtests, was administered to 830 male and female recruits to the Israel Army. The correlations among the subtests were found to be in the following Table 1:

Table 1

Correlations among the Five Subtests of

Raven's Progressive Matrices

For 830 Male and Female Recruits to the

Israel Defence Army

Subtest	A	B	C	D	E
A	1.00	.57	.53	.47	.32
B	.57	1.00	.64	.54	.35
C	.53	.64	1.00	.66	.39
D	.47	.54	.66	1.00	.52
E	.32	.35	.39	.52	1.00

These correlations show the typical order pattern of the simplex. The largest correlations are next to the main diagonal, and taper off towards the northeast and southwest corners of the Table. Such tapering alone, however, is not sufficient to prove the existence of a simplex. The next step is to compute the matric inverse of the correlation matrix in Table 1, shown in Table 2.

Table 2

The Inverse Correlation Matrix of

That in Table 1

Subtest	A	B	C	D	E
A	1.62	-.57	-.34	-.19	-.09
B	-.57	1.97	-.75	-.26	-.09
C	-.34	-.75	2.26	-.89	-.05
D	-.19	-.26	-.89	2.11	-.56
E	-.09	-.09	-.05	-.59	1.38

In the perfect simplex, the inverted correlation matrix must always show positive numbers (greater than unity) in the main diagonal, negative numbers in each diagonal adjacent to the main one, and zeros everywhere else. In practice, one expects only quasi-simplexes and not perfect ones. In practice, then, in place of zeros in all cells that are two or more steps removed from the main diagonal, one should expect a rapid tapering off toward zero. This is what Table 2 shows.

The third step in studying the structure of the data is to compute the multiple regression weights for predicting each variate from the remaining ones in the set, with results as shown in Table 3.

Table 3

Multiple Regression Weights for Predicting Each Subtest

in Table 1 from the Four Remaining Ones

Subtest Criterion	Subtest Predictor A	B	C	D	E	Multiple Correlation Coefficient
A	-	.35	.21	.12	.05	.619
B	.29	-	.38	.13	.04	.700
C	.15	.33	-	.39	.02	.747
D	.09	.12	.42	-	.28	.725
E	.06	.06	.04	.43	-	.526

The entries of Table 3 are obtained from the corresponding ones of Table 2 merely by dividing each entry of Table 2 by the main diagonal element of the same row, and reversing the algebraic sign. The main diagonal of Table 3 is left blank, since a variable is not used to predict itself. The multiple correlation coefficient for the jth variable is computed by the formula:

$$\sqrt{1 - 1/d_j}$$

where d_j is the jth diagonal element of Table 2 (or of the inverse matrix), and the results have been recorded in the last column of Table 3.

Table 3 shows directly the economy inherent in the predictive powers of a simplex. In a perfect simplex, all regressions weights would be zero except for two variates: the one immediately preceding the variate to be predicted and the one immediately following this variate. Four (or more) predictors are no better than the best two, and the best two are always the closest neighbors to the variate in question. For example, subtest C is essentially as predictable from only B and D as it is from A, B, D, and E all together.

Perhaps more striking is the fact that subtest E is as predictable from D alone as from A, B, C, and D all together. An end test in a simplex can have but one best neighbor instead of two, and its simple correlation with this neighbor must be essentially equal to its multiple correlation on all the remaining variables in the simplex. Subtest E here correlates

.52 with its best neighbor D, according to Table 1. It has substantial and significant correlations with the remaining three subtests, but the latter add nothing to the multiple regression (last line of Table 3) in addition to D; the multiple correlation is .526, which does not differ significantly from the simple correlation .52. Similarly, the first subtest A correlates .57 with its closest neighbor B, and has a multiple correlation of only .619 (Table 3) on B, C, D, and E combined.

If the simplex contained a million variates and not just five, the same parsimonious regressions would hold. Its two closest neighbors alone will give essentially as good a prediction of a particular member of the simplex as will all the million (minus the criterion) combined.

The same parsimony will tend to hold true with respect to predicting a criterion <u>outside</u> the simplex. Barring certain exceptional circumstances, an outside criterion will usually tend to have its regression weights on the variates of a simplex follow the same law of neighboring, since these weights also depend largely on the inverse of the simplex correlation matrix as well as on the correlations of the criterion with the simplex variates separately.

4a. The conclusions of the Israel Defence Army, therefore, are that for most efficient use, the Raven's subtests should be scored and used separately. If their total score should be used instead, this must necessarily correlate very highly with the middle subtest C, and in general will predict no criterion better than will subtest C alone. That is, if only a single all-purpose score must be used, subtest C alone is in general just as good as the total score on all five subtests, and of course, is just one-fifth as difficult to administer. Time and money can be saved by using only subtest C in practice, with no essential loss of predictability, if only one overall score is to be obtained from the five subtests.

(That the sum or average of the variates of a simplex must correlate highly with one of the variates of intermediate order is a general theorem. The Israel Army did not make this calculation, but from Table 1 it can be computed that the correlation of subtest C with the total is .83. The simple correlations of subtest C with outside criteria can be expected to resemble even more closely those of the total score with the selfsame criteria than this .83 might indicate by itself, for the pattern of almost-zero entries in Table 2 also militates to bring about such a similarity.)

4b. A second conclusion is that, while subtest C by itself is as good in general for predictive purposes as is the sum of all five subtests, nevertheless for particular purposes one or two of the other subtests may be even better than either C or the total score. For example, if a criterion requires a low level of complexity of the ability studied by these tests, then it will correlate highest with subtest A, and will have successively decreasing correlations with the remaining subtests. This can be checked empirically for any given criterion. Then subtest A alone will yield

essentially the same correlation with the criterion as will the multiple regression on all five subtests, and this is in general a higher correlation than can be given by the total score on all subtests. Similarly, some other criterion may require a degree of complexity between subtests D and E. Then the multiple on these two tests alone will be as efficient as on all five tests, and better than the total score. Thus less testing, using the one or two most appropriate subtests, can yield better results than using all five subtests and merely adding up the five scores.

The Raven's Test is a very popular one, being used widely throughout the world but much of its predictive power has been wasted in the past by not recognizing and making use of its simplex character.

5. In the existing literature, numerous tables of intercorrelations among many kinds of variables are reported. Many of these are not to be hypothesized to form simplexes or circumplexes or any other simple order system because of the hodge-podge nature of their content. Our project first had to examine many lists of variables which had been intercorrelated by other investigators, to judge whether enough variables could be selected from each given published set which would deal with essentially the same kind of abilities.

It was found that for the purposes of simplex analysis, three varieties of abilities were sufficiently tested in a systematic manner so as to enable us to rework published results. These three varieties are: verbal, numerical, and visual. Eight different simplexes are listed below for verbal abilities, some of them overlapping somewhat and others going off into quite different tangents. Six simplexes are reported below for numerical abilities, again some overlapping and others not. Eight simplexes of visual abilities are also reported, some overlapping and others not.

In the field of personality rating, we have selected four simplexes to be exhibited in this report, one based on the Rorschach ink blots, and the three others on types of ratings by observers.

The 26 simplexes are listed here according to source reference, the number in parentheses following each test rating being the number of the test or rating in the original reference. Each simplex listed here had the inverse matrix and multiple regressions weights calculated, as described for the example of the Raven's Test above, and the numerical results do not depart substantially from that specified by simplex theory. A far larger number of attempted examples is not reported here. Apart from imperfections in the tests themselves (that they may not be scales, for instance), a cause for aberrations from an expected simplex structure is a wandering off into neighboring but different kinds of content. This could not be explored well with the available data, except for certain circumplexes which will be reported on below.

APPROXIMATE SIMPLEXES

I. Abilities

A. Verbal Ability

1) Source: C. C. Brigham, A Study of Error, College Entrance Examinations Board, New York, 1932

The tests in the simplex order of complexity:

a - Spelling (A)

b - Punctuation (C)

c - Grammar (B)

d - Vocabulary (D)

e - Literature (E)

f - Foreign Literature (H)

2) Source: L. L. Thurstone, Factorial Studies of Intelligence, Psychometric Monographs 2. Univ. Chicago Press, 1941

The tests in the simplex order of complexity:

Parts of words (open ended questions):

a - Prefixes (42)

b - Suffixes (54)

c - First Letters (24)

d - First and Last Letters (23)

e - Four Letter Words (24)

Understanding single words

a - Same or Opposite (51)

b - Vocabulary (45)

c - Completion (10)

d - Association (6)

e - Synonyms (55)

Understanding combination of words:

a - Same or Opposite (51)

b - Vocabulary (45)

c - Sentences (46)

d - Paragraphs A (44)

Increasing abstractness of verbalization:

a - Word Checking (58)

b - Verbal Enumeration (57)

c - Same or Opposite (51)

d - Proverbs (43)

e - Reasoning (49)

Abstractness of verbalization (in another direction):

a - Proverbs (43)

b - Vocabulary (45)

c - Word Checking (58)

d - Verbal Enumeration (57)

e - Association (6)

f - Synonyms (55)

3) Source: L. L. Thurstone, Experimental Study of Simple Structure, Psychometric, Vol. 5 No. 2 June, 1940

The test in the simplex order of complexity:

Individual Letters of the Alphabet:

a - Scattered X's (29)

b - Repeated Letters (27)

c - Letter Grouping (13)

d - Letter Series (14)

e - Word Patterns (35)

Letters to thoughts:

a - Letter Grouping (13)

b - Letter Series (14)

c - Word Patterns (35)

d - Proverbs (22)

e - Same or Opposites (28)

f - Completion (5)

B. Numerical Ability

1) Source: L. L. Thurstone, Primary Mental Abilities, Psycnometric Monographs No. 1, Univ. Chicago Press, 1938

The tests in the simplex order of complexity:

a - Addition (31)

b - Substraction (32)

c - Multiplication (33)

d - Division (34)

e - Arithmetical Reasoning (39)

f - Numerical Judgment (38)

2) Source: L. L. Thurstone, Factorial Studies of Intelligence, Psychometric Monographs 2. Univ. Chicago Press, 1941

The tests in the simplex order of complexity:

a - Number Patters (38)

b - Identical Numbers (30)

c - Multiplication (37)

d - Addition (3)

e - Arithmetic (5)

3) Source: L. L. Thurstone, Experimental Study of Simple Structure, Psychometrika Vol. 5. No. 2 June, 1940 (p. 163)

The tests in the simplex order of complexity:

a - Addition (1)

b - Multiplication (18)

c - Arithmetic (3)

d - Number Series (20)

e - Squares (31)

4) Source: P. J. Olckers, "A Factorial Study of Arithmetical Ability," Journal of Social Research, Pretoria, South Africa Vol. 2. No. 1 June, 1951

The tests in the simplex order of complexity:

From Table 7, page 4:

a - Dot Counting (6)

b - Multiplication (3)

c - Subtraction (2)

d - Division (4)

From Table 4, page 9:

a - Four Rules A (1)

b - Four Rules C (6)

c - Fractions (4)

d - Number Series (11)

e - Problems (10)

From Table 7, page 14:

a - Four Rules A (3)

b - Three Digit Addition (12)

c - Four Rules B (8)

d - Fractions (6)

e - Arithmetic (21)

C. Visual Abilities

1) Source: L. L. Thurstone, Factorial Studies of Intelligence, Psychometric Monographs 2. Univ. Chicago Press, 1941

The tests in the simplex order of complexity:

a - Dot counting I. (14)

b - Dot counting III. (16)

c - Dot counting II. (15)

d - Pursuit (44)

e - Mazes I. (35)

f - Mazes II. (36)

2) Source: A. Lubin and A. Summerfield, "A square root method of selecting a minimum set of variables in multiple regression," Psychometrika, Vol. 16. No. 4. Dec. 1951

The tests in the simplex order of complexity:

a - Stenquist Picture I. (3)

b - Minnesota Assembly (4)

c - Minnesota Spatial Relations (1)

d - Paper Form Boards (2)

e - Interest Blank (5)

3) Source: L. L. Thurstone, Primary Mental Abilities, Psychometric Monographs No. 1. Univ. Chicago Press, 1938

The tests in the simplex order of complexity:

First Set:

a - Lozenges A (19)

b - Surface Development (23)

c - Pursuit (27)

d - Areas (29)

e - Hands (53)

Second Set:

a - Figure Classification (8)

b - Pattern Analogies (44)

c - Block Counting (14)

d - Identical Forms (26)

Third Set:

a - Figure Classification (8)

b - Copying (28)

c - Form Board (21)

d - Block Counting (14)

Fourth Set:

a - Cubes (18)

b - Flags (20)

c - Lozenges B (22)

d - Punched Holes (24)

Fifth Set:

a - Pursuit (27)

b - Figure Classification (8)

c - Copying (28)

d - Identical Forms (26)

Sixth Set:

a - Lozenges B (22)

b - Punched Holes (24)

c - Form Board (21)

d - Areas (29)

II. Personality Traits

1) Source: L. L. Thurstone, A Factorial Study of Perception, Univ. Chicago Press, 1944

The five Rorschach scores in their simplex order of complexity:

a - Weight Color

b - Perceptual Organization

c - Movement

d - Total Responses

e - Total Time

2) Source: W. M. O'Neil, "The Construction of a Staff Report Form," Occupational Psychology, Vol. XXVI. No. 3. July, 1952

The ratings in their simplex order of complexity:

a - Accuracy

b - Speed

c - Application

d - Personality

3) Source: The OSS Assessment Staff, Assessment of Men, Reinhart Co., New York, 1948

The Social Relations Ratings in simplex order of complexity (from Table 62, page 520):

a - Brook

b - Assigned Leadership

c - Interview

d - Discussion

The Energy and Initiative Ratings in simplex order of complexity (from Table 59, page 517):

a - Obstacle

b - Brook

c - Assigned Leadership

d - Debate

e - Discussion

6. From the above list of simplexes, it will be clear that some of the original conjectures of the monograph in press on the radex must already be taken into account. For instance, the first example discovered of a simplex since the beginning of radex theory was of certain verbal ability tests. The nine tests and their intercorrelations were taken from the Thurstones' Factorial Studies of Intelligence. The inverse matrix did show somewhat aberrant tapering to zero. Closer examination has now broken this set up into parts that go with other subsets of tests into various different — albeit intercorrelated — simplexes in the above list. This part of the radex structure, therefore, by itself indicates that the initial two-dimensional radex of the monograph already is inadequate to portray the intricacies of the functional relationships. Higher dimensional radexes must be used to portray all the tests simultaneously. This is further verified by the fact that a circumplex was found among certain verbal tests alone, as will be reported below in discussing findings on the circumplex.

It is difficult to make a deeper psychological analysis of the structure of verbal abilities here, because the most interesting data to compare are from different studies and hence do not have their intercorrelations available. It is sufficient to summarize at this stage that there are several possible different, albeit overlapping, chains of increasing complexity for verbal understanding, as from "Same or Opposite" to "Paragraphs", "Word checking" to "Reasoning" and "Proverbs" to "Synonyms". Similarly, there are several different pathways of increasing complexity from parts of words to complete thoughts, as from "Letter grouping" to "Completion" and "Spelling" to "Foreign Literature".

Numerical abilities generally follow the expected path of complexity from addition through division. However, an important point emerges. Addition is more complex than multiplication in one simplex, where six two-digit numbers were to be added while a two-digit number was to be multiplied by only a one-digit number. Complexity here then depends on the sheer number of digits as well as on the nature of the operations.

The visual ability simplexes are usually quite different from each other in the types of tests involved, but generally show that what looks more complicated in existing tests does require more complex behavior. Also, as might be expected, visualizing solids is more complex than plane figures, than linear ones. Different possible chains of complexity are possible here too, involving some of the same tests, as "Lozenges" and "Punched Holes".

7. It was not expected that personality traits would in general form ordered systems of the same kind as do mental abilities. A major reason for this is the growing pile of evidence at the Israel Institute of Applied Social Research that polytonicity and multivaluedness very frequently characterize the regression systems of such data. However, in some cases linear least squares regressions may hold, and hence a radex

in the present sense may be possible. On the other hand, the appearance of such a simple regression system may be an artifact of the scoring system used, or of ignoring basic features of the phenomena being studied.

In any event, certain Rorschach scores used in a study by Thurstone were found to form an approximate simplex, as listed above. Of the many ratings intercorrelated by the OSS staff, two subsets provided clear simplexes. Many others seem to be fairly approximate simplexes, but the samples were often too small to be reliable, and not enough additional variables were present to study aberrations.

8. One of the most gratifying results of the project is the verification of the existence of a basic circular ordering for certain abilities that differ in <u>kind</u> and not just in degree of complexity. The circle is, of course, divisible into many segments, or one ability shades into the next gradually, but the following gross landmarks may serve for purposes of orientation. The circle may be regarded as going from verbal abilities, through numerical abilities, then visual, then abstract thinking, and then back to verbal. This is illustrated by the following table of correlations.

<u>Table 4</u>

An Approximate Circumplex of Various Abilities for

710 Chicago Schoolchildren*

Test	Association (6)	Incomplete Words (32)	Multiplication (37)	Dot Patterns (17)	ABC (1)	Directions (12)
(6)	1.	.446	.321	.213	.234	.442
(32)	.446	1.	.388	.313	.208	.330
(37)	.321	.388	1.	.396	.325	.328
(17)	.213	.313	.396	1.	.352	.247
(1)	.234	.208	.325	.352	1.	.347
(12)	.442	.330	.328	.247	.347	1.

* Data from L. L. Thurstone, <u>Factorial Studies of Intelligence</u>, Psychometric Monographs 2. Univ. Chicago Press. 1941.

The inverse correlation matrix was computed from Table 4, and then the multiple regression weights were obtained for predicting each test from all the remaining ones, as shown in Table 5.

Table 5

Multiple Regression Weights for Predicting

Each Test in Table 4 from All the Rest

	Predictor						Multiple Correlation Coefficients
Criterion	Association (6)	Incomplete Words (32)	Multiplication (37)	Dot Patterns (17)	ABC (1)	Directions (12)	
(6)	—	.30	.10	-.01	.04	.30	.553
(32)	.31	—	.20	.15	-.02	.09	.540
(37)	.10	.21	—	.23	.14	.11	.534
(17)	-.01	.16	.25	—	.23	.04	.486
(1)	.04	-.02	.15	.23	—	.23	.465
(12)	.31	.10	.11	.03	.21	—	.533

The circular order existing among the tests in Table 4 is reflected by the cyclic rise and fall of the correlation coefficients in each row (column). In contrast to the previous case of the simplex, the "last" test in Table 4 is an immediate neighbor of the "first" one. Which test is called "first" is arbitrary in a circumplex, since the tests differ in kind and not in degree. The circular order is more precisely established by Table 5. In a perfect circumplex, all regression weights must be zero for any test not an immediate neighbor of the one being predicted. In an imperfect or quasi-circumplex non-neighbors should tend to have zero or relatively small weights. Thus, in Table 5, "Association" is about as predictable from only its two immediate neighbors, "Incomplete Words" and "Direction", as it is when the three remaining tests are added as predictors.

Although different tests were involved, the same circle of verbal-numerical-visual-reasoning-verbal was found for the same children as of Table 4 when greater complexity was involved on the average in each test. The two circumplexes plotted together give already a rough radex with tests differing in kind as well as in degree in the same picture. Another battery of Thurstone's administered to a different group of children reveals the same kind of circle. The Bellevue-Wechsler battery also reveals the same circle for two different age groups of adults.

The well-tested circumplexes of this project are as in the following list.

APPROPRIATE CIRCUMPLEXES

Each set is arranged in circumplex order. It is, of course, arbitrary which test is listed first in the set; another choice would mean merely a cyclic rearrangement of the given l''ting.

I. Abilities

1) Source: L. L. Thurstone. Factorial Studies of Intelligence. Psychometric Monographs 2. Univ. Chicago Press, 1941

Several Different Abilities (one level of complexity):

a - Association (6)

b - Incomplete Words (32)

c - Multiplication (34)

d - Dot Patterns (14)

e - ABC

f - Directions

Several Different Abilities (another level of complexity):

a - Digit Span (11)

b - Arithmetic (5)

c - Geometrical Forms (28)

d - Identical Pictures (31)

e - Picture Naming (41)

f - Rhyming Words (50)

Different Verbal Abilities:

a - Rhyming Words (50)

b - Completion (10)

c - Association (6)

d - First Letters (24)

2) Source: D. Wechsler, The Measurement of Adult Intelligence, Third Edition, Baltimore, 1944

For Adults, Ages 20-34 (page 223):

a - Digit Span (3)

b - Arithmetic (4)

c - Block Designs (7)

d - Picture Completion (6)

e - Comprehension (1)

For Adults, Ages 35-49 (page 224):

a - Digit Span (9)

b - Arithmetic (4)

c - Block Designs (4)

d - Picture Completion (6)

e - Comprehension (1)

3) Source: L. L. Thurstone, Primary Mental Abilities, Psychometric Monographs, No. 1, Univ. Chicago Press, 1938

Several Different Abilities (one level of complexity):

a - Division (34)

b - Pursuit (27)

c - Block Counting (17)

d - Mechanical Movement (25)

e - False Premises (42)

f - Theme (52)

g - Sound Grouping (55)

Several Different Abilities (another level of complexity):

a - Number-Number (48)

b - Addition (31)

c - Pursuit (27)

d - Block Counting (17)

e - Completion (11)

f - Word Number (46)

Different Verbal Abilities:

a - Pattern Analogies (44)

b - Code Words (43)

c - Verbal Analogies (41)

d - Controlled Association (9)

e - Completion (11)

Different Verbal Abilities (a circle overlapping the previous one):

a - Pattern Analogies (44)

b - Code Words (43)

c - Verbal Analogies (41)

d - Reasoning (40)

e - False Premises (42)

4) Source: L. L. Thurstone, "Experimental Study of Simple Structure," Psychometrika, Vol. 5. No. 2. June, 1940, (page 163)

A Circumplex formed by omitting the least complex numerical ability (addition) from a short simplex, and adding a test (number patterns) that closes a circle:

a - Multiplication (38)

b - Arithmetic (3)

c - Number Series (20)

d - Squares (31)

e - Number Patterns (19)

5) Source: P. J. Olckers, "A Factorial Study of Arithmetical Ability," Journal of Social Research, Pretoria, South Africa Vol. 2. No. 1. June, 1951, Table 4 (page 9)

   a - Four Rules A (1)

   b - Four Rules B (4)

   c - Fractions (7)

   d - Change (8)

   e - Number Series (11)

6) Source: Clyde Coombs, "A Factorial Study of Numerical Ability," Psychometrika, Vol. 6. No. 3. June, 1951, Table 3

   a - Figures (26)

   b - Size Comparison (15)

   c - A B (4)

   d - A B C (5)

   e - Forms (6)

   f - Marks (30)

7) Source: G. H. Thomson, An Analysis of Performance Test Scores for a Representative Group of Scottish Children, Univ. of London Press, 1940

   a - Binet I. Q.

   b - Kohs Block Design

   c - Cube Construction

   d - Red Riding Hood

   e - Healy Picture Completion

8) Source: Ch. H. Goodman, "A Factorial Analysis of Thrustone's Sixteen Primary Mental Abilities," Pschometrika, Vol. 8, No. 3. September, 1943

a - Number Series (15)

b - Cards (7)

c - Identical Forms (1)

d - Verbal Enumeration (2)

e - Same or Opposite (6)

f - Completion (5)

II. Personality Traits

1) Source: F. T. Tyler, "A Factorial Analysis of Fifteen MMPI Scales," Journal of Consulting Psychology, Vol. 15. No. 6. December, 1951

First Set:

a - Psychopathic Deviate

b - Schizophrenic

c - Prejudice

d - Social Introversion

e - Depression

f - Hypochondriasis

Second Set:

a - Masculinity - Feminity

b - Status

c - Dominance

d - Paranoia

e - Hysteria

f - Responsibility

g - Hypomania

2) Source: W. M. O'Neil, "The Construction of a Staff Report Form," Occupational Psychology, Vol. XXVII. No. 3. July, 1952

a - Mental Alert (6)

b - Knowledge (7)

c - Accuracy (1)

d - Speed (2)

e - Applications (3)

3) Source: Ann Roe: "Psychological Tests of Research Scientists" Journal of Consulting Psychology, Vol. 15. No. 6. Table 3, December, 1951

a - Verbal, number tried

b - Spatial, number tried

c - Mathematical, number tried

d - Rorschach

e - TAT length

4) Source: The OSS Assessment Staff, Assessment of Men, Reinhart Co., New York, 1948

The Leadership Ratings (from Table 63, page 521):

a - Interview

b - Brook

c - Discussion

d - Debate

e - Ratings by Associates

The Propaganda Skills Ratings (from Table 67, page 524):

a - O W I

b - Manchuria

c - Interview

d - Discussion

9. As the above list shows, a circumplex was found within certain verbal abilities alone, as well as within certain numerical abilities alone and within certain visual abilities alone. This proves the possible existence of a continuum of circularly ordered abilities within each kind, as well as a general circular ordering between the kinds. For this reason, one of the purposes of the project — to provide a simple two-dimensional radex map of all the abilities simultaneously — could not be carried out. A far more comprehensive new testing program, directly aimed at this problem, is needed to provide the basic correlation matrix necessary, and a two-dimensional map will certainly not be adequate, although a three-dimensional sphere may prove to be.

This failure, however, does not negate the practicality of the present findings. Better predictions can be obtained from the sets of tests as is, and with less work, by capitalizing on their circumplex properties alone (regardless of their ultimate place in a more complex radex) than by current usage.

10. With respect to personality traits, it may be interesting to find in our list of circumplexes two sets of variables from the Minnesota Multiphasic Personality Inventory. Certain personnel ratings also have been found to form circumplexes, as listed above. As remarked in conjunction with our simplexes, this may be due in part to the scoring techniques, since multi-valued and polytonic relations may be more typical of such data.

Results of the Project with Respect to Multivalued and Polytone Attitudinal Interrelations

1. Data from three field studies of the Israel Institute of Applied Social Research were found somewhat appropriate for the new kind of analysis developed in this project for multivalued and polytone relations. Ideal data for this purpose proved difficult to find, since the underlying theory was developed after the completion of some large surveys conducted by the Institute, and no new studies had been made to the time of this project which could be used for this methodological end. Nevertheless our results are sufficiently positive to substantiate the major hypothesis. The future will undoubtedly reveal neater data for further verification, since we now have learned more of what is needed for more clearcut results.

2. A typical table (in many respects) for this project is shown below as Table 6.

Table 6

Frequency Cross Tabulation of Two Scalable Attitudes

Hypothesized to Have Monotonely Related Involutions

Rank on "Mixed Housing"	Rank on "Moroccans as Neighbors" 1	2	3-4	5	6	7	8	9	Total
0	1	2		56	1		45	9	114
1	4	3		79	2	2	75	11	176
2	1	8	2	150	4	4	105	10	284
3	4	14	1	73	3	7	126	13	241
4		1	3	74	6	6	67	10	167
5	1	1	2	24	4	2	26	4	64
6	1	2	3	105	11	3	56	15	196
7-8	1	3		46	3	8	73	19	153
9-10	4	6	1	40	5	10	57	30	153
Total	17	40	12	647	39	42	630	121	1548

The horizontal variable is ranks on a quasi-scale of attitude toward living together with Moroccans in a housing project. The vertical variable is ranks on a scale of attitude toward living in a "mixed" housing project, that is, one containing immigrants from various countries of origin — but not specifying the countries. The analysis of the table is based on the following considerations.

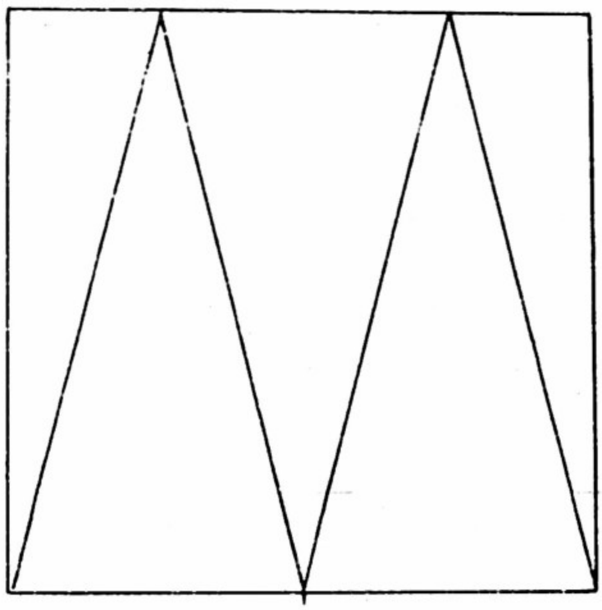
Figure 1

Theoretical Relationship of the Involution

(Fourth Principal Component) of a Scalable

Attitude With that Attitude

Involution



Attitude

The psychological theory of principal components of scales has been developed elsewhere (cf. Louis Guttman, "The Principal Components of Scalable Attitudes," in Paul F. Lazarsfeld editor, Mathematical Thinking in the Social Sciences, in press). According to this the involution, or fourth principal component of a scale like attitude toward Moroccans should plausibly be a scale on: "Is the respondent actively reviewing his position on this problem or not?" Since the data on hand were assembled before the psychological theory of the fourth principal component was clear, a scale directly aimed at the involution of "attitude toward Moroccans as neighbors" was not included. Had such an involution scale been available, it should have yielded ideally a perfect curvilinear regression on the attitude toward Moroccans of the type shown in Figure 1.

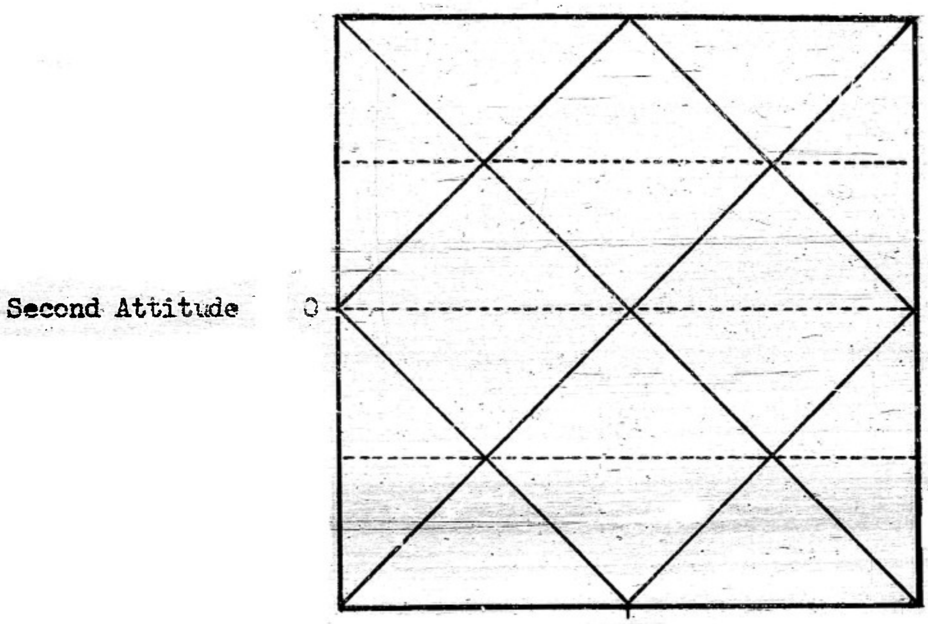
Now, suppose that a second attitude had exactly this same involution for its own fourth principal component. In such a case, we must have the striking type of joint relationship between the two attitudes illustrated by Figure 2.

Figure 2

Theoretical Relationship Between

Two Different Scalable Attitudes

Having the Same Involution



First Attitude

[In both Figures 1 and 2, we have drawn the regression curves as symmetrical about the zero point (as defined by the intensity or second principal component) and also as polygon rather than with smooth bends. This is only for convenience here, and does not of course represent the general case.]

The novel feature of Figure 2 is that it shows each attitude to be in general a <u>multivalued function</u> of the other. To each value of one attitude, there corresponds in general more than one ideal value of the other. This is a result of the <u>polytone</u> regression of the common involution on each attitude separately, as in Figure 1. In Figure 1, the attitude continuum can be broken up into four segments in each of which there is a <u>monotone</u> relation of the involution to the attitude, the <u>direction</u> or <u>sign of the relationship</u> alternating from segment to segment. The same would be true of the corresponding figure for the same attitude. In cross-tabulating the two attitudes to get Figure 2, it must be that each monotone

segment of the second attitude must allow the full polytone relationship to appear with respect to the first attitude. This is shown in the top segment, cut off by the first inserted broken line, in Figure 2. The next segment must allow the same bendings, but with reversed direction, etc. The same holds if we regard in turn monotone segments of the first attitude and how they relate to the second.

Now, it is not hypothesized that the two empirical attitudes in Table 6 above have the same involution. To the contrary. Instead, it is hypothesized that the two involutions have a monotone relationship with each other in each direction. The more a respondent is actively reviewing his position on Moroccans as neighbors, the more -- on the average -- he is reviewing his position on mixed housing in general; and vice versa. If this be true, Table 6 should resemble Figure 2, but with random error -- indeed considerable error. Our calculations tend to verify this.

3. The quantile technique, as illustrated for Table 6, consists simply of computing for each column or crude scale-rank, as many position values as rows along the percentile metric of the vertical scale, and then doing the analogous computing for each row. For example, the row for ranks 9-10 in Table 6 has a marginal total of 153 out of the grand total of 1548, or 9.8%. This means this rank covers the percentile from 0 to 9.8, with midpoint at 4.9. In each rank column separately, therefore the 9.8 percentile is sought, and then interpolated into the percentile metric of the marginal total, and again the midpoint taken. Thus the 9.8 percentile from the bottom of column 9 of the table is $.098 \div \frac{30}{121}$ of the way up the row 9-10, or .39 of the interval from zero to $\frac{153}{1548}$ of the marginal total. The midpoint of this is plotted as in Figure 3.

Each column has as many quantiles as there are rows, and the corresponding ones are connected by straight lines. If there were complete statistical independence, the lines would all be perfectly straight and horizontal. If there were a monotone regression of the vertical scale on the horizontal one, the lines would all tend to slope in one direction, say from the southwest to the northeast corner of the picture.

But we see a polytone pattern revealed by Figure 3, roughly as hypothesized.

Computing the quantiles in the analogous manner, but for each row of Table 6 leads to the picture in Figure 4. Again a polytone pattern is revealed.

Figure 3

The Quantiles for Each Column of Table 6

In the Percentile Metric of the Row Marginal Totals

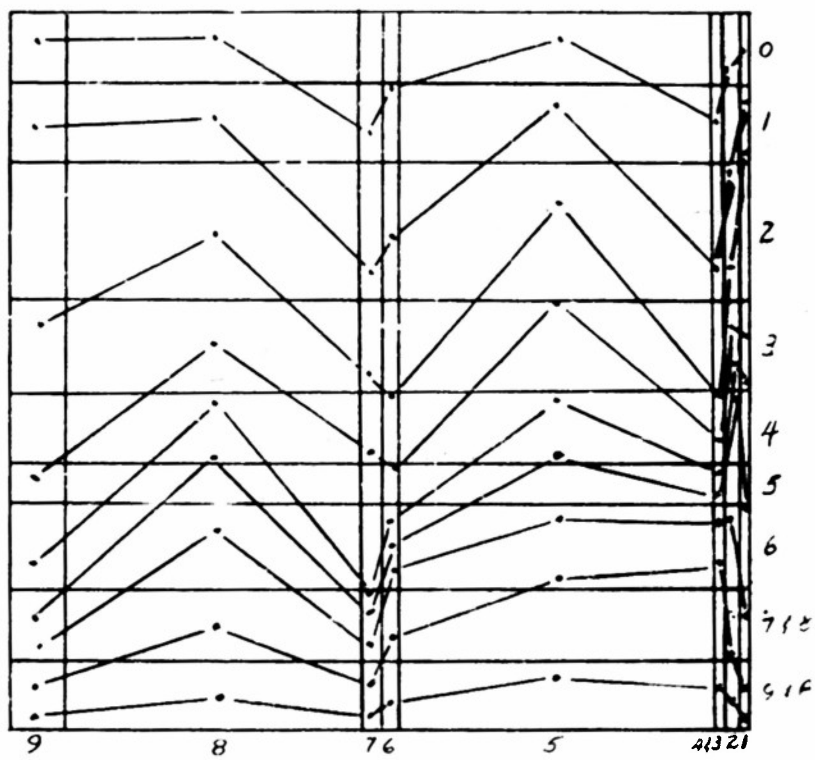
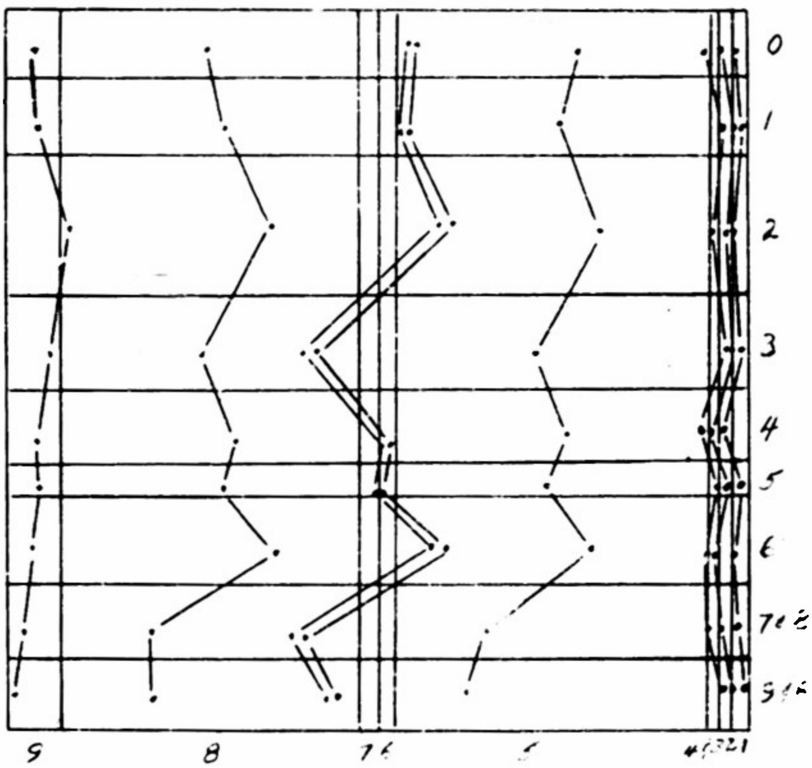


Figure 4

The Quantiles for Each Row of Table 6

In the Percentile Metric of the Column Marginal Totals



4. The drawbacks of the quantile procedure are quite evident here, as well as its virtues. Both Figure 3 and Figure 4 should resemble Figure 2, taking error into account. There are at least two sources of error. One is that the (hypothesized) relationships between the two involutions is not perfect, albeit monotone; there is regression error hypothesized in both directions. The other is the usual one of having only approximate scales to begin with and not perfect ones with some resultant unreliability of scale ranks.

The quantile computations can introduce a third source of error, namely the arbitrariness of the observed rank interval, or of the observed cutting points along the two scales. Even if the first two sources of error were not present, this third source could still distort the quantile pictures. The third source creates disturbance whenever a bending point of the common principal component -- as in Figure 1 -- does not coincide exactly with an observed cutting point point of a scale. This results in a scale rank or interval which straddles a bending point and hence does not permit a monotone relation between scale and component within this interval.

In Table 6, it appears that substantial straddling may be occurring, leading to distorted results in regions where bending points may belong. More than three bends definitely occur in some levels of the quantiles; this can be explained by such straddling alone. That the curves in Figure 3 (or Figure 4) do not clearly alternate in direction in four different regions, as in Figure 2, can also be explained by this straddling, especially if one or two of the monotone regions as in Figure 2 have small marginal frequencies. They would then be dwarfed in the computing by their neighboring sections and appear almost not to exist.

A fourth source of error is the usual one due to size of sample. This is effective especially in rows (columns) with the smallest marginal totals, when the first two types of error are also present. This is also in evidence in Figures 3 and 4.

5. The latent vector technique of analysis avoids some of the drawbacks of the quantile technique by not being entirely dependent on fluctuations of the data within small local regions of a table such as Table 6. It analyzes the table as an interdependent whole, and not each part (row or column) separately and then interrelating the parts. Its general logic is as follows:

Given a table of joint frequencies between two variates X and Y. Let P_{jk} be the proportion in row j and column k of the table. Let the row and column marginals be, respectively:

$$P_{j.} = \sum_k P_{jk} \qquad P_{.k} = \sum_j P_{jk}$$

and let N be the total number of cases in the Table:

$$N = \sum_j P_{j.} = \sum_k P_{.k} = \sum_{jk}\sum P_{jk}.$$

The respective ranges of j and k can be quite different; i.e., the Table need not be square. x and y can differ in the number of categories each has.

Let us seek a real number for each row of the Table, say x_j, and a real number for each column, say y_k, such that the two resulting sets of numbers will have as linear a relation as possible in each direction, and with a maximum product-moment correlation coefficient. There is no loss of generality if we assume the marginal means to be zero:

$$(1) \qquad \sum_j x_j p_{j.} = \sum_k y_k p_{.k} = 0$$

Then the correlation to be maximized is:

$$(2) \qquad r = \frac{\sum\sum_{jk} x_j y_k p_{jk}}{\sqrt{(\sum_j x_j^2 p_{j.})\ (\sum_k y_k^2 p_{.k})}} .$$

The maximizing equations can be obtained by differentiating (2) directly, or by using the method of Lagrange multipliers, as follows. Let ϕ be defined as

$$(3) \qquad \phi = \sum\sum_{jk} x_j y_k p_{jk} - \tfrac{1}{2}\lambda \sum_j x_j^2 p_{j.} - \tfrac{1}{2}\mu \sum_k y_k^2 p_{.k} ,$$

where λ and μ are Lagrange multipliers. Differentiating (3) with respect to the y_k and setting the results equal to zero yields:

$$(4) \qquad \sum_j x_j p_{jk} = \mu\, y_k p_{.k} \quad .$$

Differentiating (3) with respect to the x_j and setting the derivatives equal to zero yields:

$$(5) \qquad \sum_k y_k p_{jk} = \lambda\, x_j p_{j.} \quad .$$

Multiplying (4) by y_k on both sides and summing over k, multiplying (5) by x_j and summing over j, and comparing with (2) shows that

$$(6) \qquad r^2 = \lambda\mu .$$

The equations (4) and (5) are of the same nature of the equations of internal consistency of scale analysis (cf. Louis Guttman, "The Quantification of a Class of Attributes," in Paul Horst, et al., The Prediction of Personal Adjustment, Social Science Research Council, 1941), although the context is quite different. As in the case of scale analysis, equations (4) and (5) lead not to one solution but to a series of solutions, called principal components or latent vectors of two certain matrices but for which in general there is but one best one. The major latent root μ is usually unique and provides the desired maximum r^2, and associated with it are the maximizing x_j and y_k or elements of the major latent vectors.

Equations (4) and (5) can be solved by iterations, passing from a trial x to a trial y and back again according to these equations, or else by first computing the matrix A with typical element a_{jk} defined by

$$(7) \qquad a_{jk} = \frac{1}{p_{j.}} \sum_l \frac{p_{jl} p_{kl}}{p_{.l}} .$$

Then the desired vector $x = (x_1\ x_2 \ldots)$ is a latent vector of A corresponding to root r^2, or in matric notation,

$$(8) \qquad xA = r^2 x.$$

It is equation (8) we used in practice in this project, with the new iterative technique referred to above. The first trial x must be chosen to have a zero mean.

Once a solution $x = (x_1\ x_2 \ldots.)$ has been arrived at, the corresponding solution $y = (y_1\ y_2 \ldots.)$ can be obtained directly from equation (4) (the constant of proportionality μ being arbitrary then, and can be chosen to be $\mu = 1$).

6. For our problem, the subscripts j and k have a definite meaning. They represent actual rank orders on two scales. If the maximizing x_j and y_k maintain the same rank order, then we are justified in stating that the two scales have a monotone regression each on the other. If the x_j change rank order but the y_k stay put, then x is a polytone function of y. Conversely, if the y_k change rank order but the x_j stay put as far as their ranks go, then the y scale is a polytone function of the x scale. If rank orders change both within the x_j and the y_k, then we have a multivalued and polytone relationship in both directions of which Figure 2 above is an example.

As an illustration of actual results, consider the major latent vectors for Table 6. For the eight columns, the y_k are (up to an arbitrary constant of proportionality, and rounded off to two digits) as in Table 7.

Table 7

Relation of Its Major Latent Vector Elements

to the Scale Ranks of "Moroccans as Neighbors"

of Table 6

Scale rank	1	2	3-4	5	6	7	8	9
Vector value	12	8	-7	-6	-2	14	2	11

Plotting the second row of Table 7, as a function of the first, as in Figure 5 below, shows how the ranks of the vector values change from those of the original scale, and in the manner predicted by our hypothesis of the fourth principal component. According to Table 7 and Figure 5, what the scale of "Moroccans as Neighbors" has most in common linearly with the other scale is its own involution. Of the attitude system toward "Moroccans," it is the aspect of involution that is most closely related to the attitude system of "Mixed Housing," and not the direct content of the attitude.

Our data show the same thing in the opposite direction. For the eight rows of Table 6, the x_j are (up to an arbitrary multiplying constant, and rounded off to two digits) as in Table 8.

Table 8

Relation of Its Major Latent Vector Elements

to the Scale Ranks of "Mixed Housing"

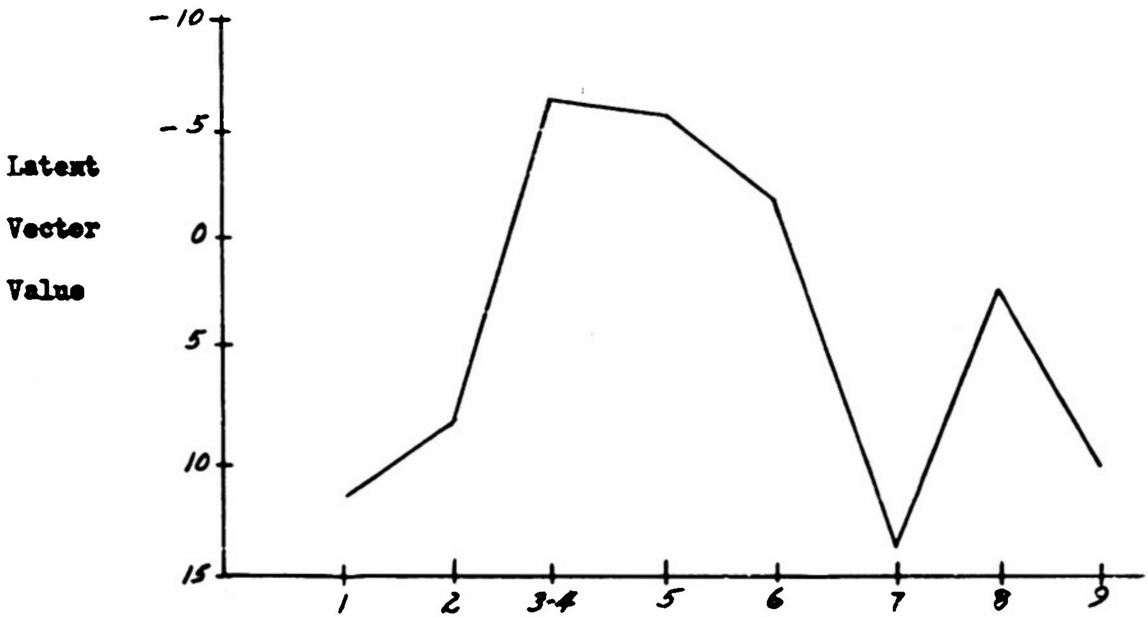
of Table 6

Scale rank	0-1	2	3	4	5	6	7-8	9-10
Vector value	-5	-14	10	-6	-2	-14	16	29

Except for a slight aberration at rank 5, the rearrangement of ranks by the vector is again according to our hypothesis of monotonely related fourth principal components between the two scales. This is shown graphically in Figure 6.

Figure 5

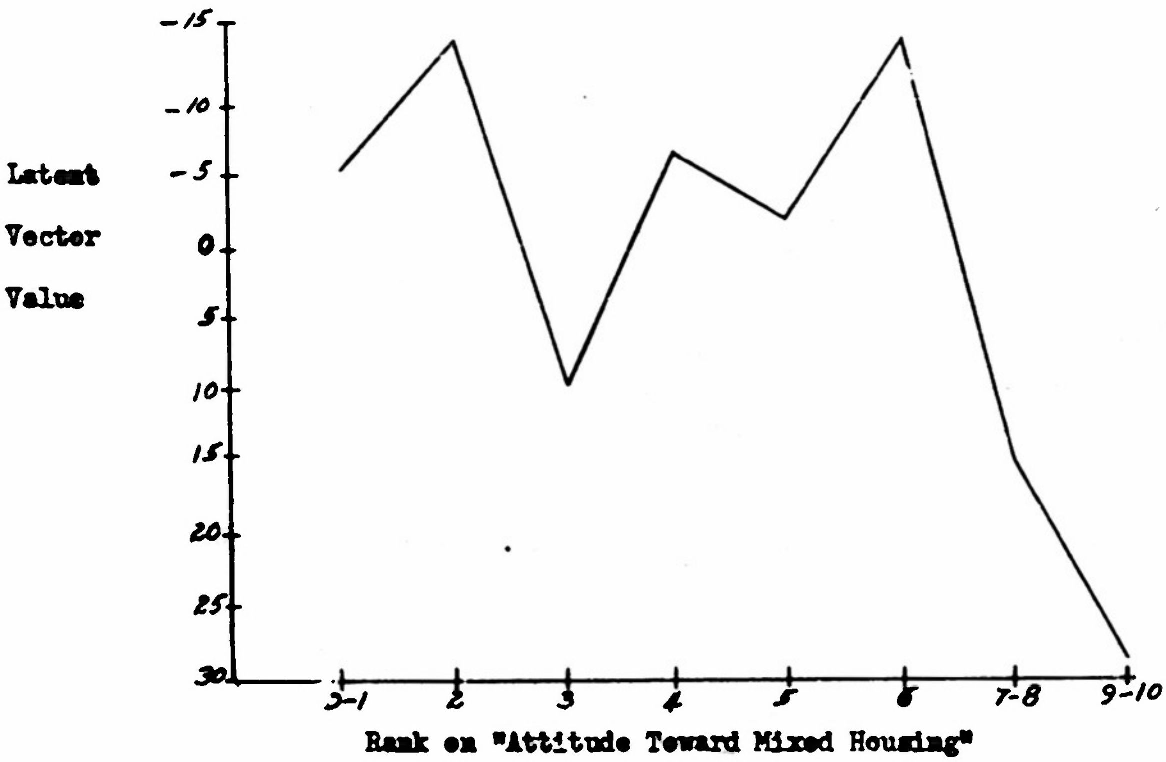
The Polygon of the Data of Table 7



Rank of "Attitude Toward Moroccans as Neighbors"

Figure 6

The Polygon of the Data of Table 8



Tables 7 and 8, therefore, imply that the fourth principal components of the scales are monotonely related to each other, and that the relationship in Table 6 between the scales themselves must be multivalued, of the nature depicted by Figure 2.

7. The latent vector technique has its drawbacks as well. We have stressed that a picture like in Figure 2 may occur if our hypothesis is correct. However, it need not always occur, even though the hypothesis is correct. Parts of the picture may be missing, or have zero frequencies. An example of this is shown in Table 9. It may be hypothesized that "Attitude toward Moroccans as Neighbors" should have its involution monotonely related to "Attitude toward Yemenites as Neighbors," and this does seem to be true. However, as Table 9 shows, the full possibilities of multivaluedness do not empirically occur.

Table 9

Cross-Tabulation of "Moroccans as Neighbors" with "Yemenites as Neighbors"

Rank on "Yemenites"	Rank on "Moroccans"								
	1	2	3	4	5	6	7	8	9
1									25
2					1			61	1
3							3		6
4						11			
5					753				
6				23		22			
7			12						39
8		139						634	
9	23		8				24		94

Many aspects of the scattergram of Table 9 are largely spurious, and that is why the Table looks so neat. The scales on "Moroccans" and "Yemenites" are not of the usual type, but were each derived in large part from two essentially open ended questions. Hence the relative lack of error in the Table. In particular, the lack of spread in either direction among

the 753 middle cases is almost completely spurious. But this spuriousness aside, Table 9 may be helpful in illustrating the possible appearances of multivaluedness. The multivaluedness itself is not spurious, although the lack of error around it is. Table 9 can be compared almost directly with Figure 2 to see how parts of the picture can occur and how other parts can be missing, even ideally.

In particular, it should be clear, only a monotone relationship may occur, even though a more multivalued one might be possible.

When parts of the picture are missing, this may throw off the latent vector technique. This technique works best when frequencies are fairly uniform over symmetrical portions of the entire ideal multivalued possibilities. Thus, the latent vector technique is not very appropriate to Table 9, since the frequencies occur symmetrically here.

8. Because of the complexity attending multivalued relations, it is not to be expected that they can be easily studied unless the data themselves are "pure" enough for the purpose. "Scales" arrived at by item analysis and other techniques in the past may not lead to any further productive analysis in the present direction since they in general have no meaningful principal components. With more care in the initial defining of attitude areas, and more careful delimiting of unidimensionality — as by scalogram and related techniques — the deeper psychological analyses revealed to be possible by the present project can be expected to become more possible and more essential in the future.

9. Even the Institute's own data from its past researches revealed many deficiencies in the light of our present knowledge, despite the care usually spent on scale analysis. The following is a list of topics from previous studies of the Institute which proved somewhat adequate for the present purposes of this project, and the results as briefly indicated.

ANNOTATED LIST OF CROSS-TABULATIONS

OF ATTITUDE SCALES

(Analysis by latent vectors and/or quantiles for the study of polytonicity and multivaluedness)

Study A: Postwar Plans of Soldiers Before Discharge

Sample: 2173 soldiers, a cross-section of the Defence Army of Israel.

Scales Cross-tabulated: Each of the content scales below refers to a particular alternative postwar plan, and all have the same third principal component (or closure) in common. The closure was related to each content separately, and each content was related to every other content.

The scales:

(a) Closure (definiteness of choice among the various alternatives)

Attitudes toward the following respective alternatives:

(b) Permanent military service

(c) Government job

(d) Police

(e) Urban cooperative enterprise

(f) Independent business

(g) Occupational training

(h) Return to previous job

(i) Further schooling

(j) Agricultural settlement

Overall Results: Both quantiles and latent vectors verified in general the polytonicity and multivaluedness predicted by the hypothesis that the closure was common to all the contents. Some minor discrepancies may be due to imperfection of scales and/or sampling error.

Study B: Social Aspects of Housing Projects

Sample: 1881 women in a sample of all new housing developments in Israel.

Scales Cross-tabulated: Each of the content scales below was cross-tabulated with every other. The involutions of each all seem to be monotonely interrelated.

The scales: Attitudes toward:

(a) Mixed housing

(b) General neighboring

(c) Yemenites as neighbors

(d) Moroccans as neighbors

Study C: Opinion Survey

Sample: 2147 adults over 18 years of age, a cross-section of all Israel residents (except Arabs).

1. Attitudes toward: equal rights for women and military service for women.

   Monotone relation in one direction, but apparently polytone in the other.

2. Attitudes toward: salary level for Knesset members and salary level for government officials.

   Monotone and univalued in one direction, but other direction has indication that attitudes toward salaries of government officials may relate monotonely to the fourth principal component of salary for Knesset members.

APPENDIX I

Original Statement of the

Purpose of the Project

The problem of factor analysis remains an important one, both for the fields of psychometrics and of attitude research. When a battery of mental tests is used, it is essential to know its structure in order to make the best use of the battery. Similarly, it is important to know the structure of the interrelationships of systems of attitudes in order to comprehend and to be able to predict human behavior. The Israel Institute of Applied Social Research has had to face the problem of factor analysis in both of these fields, that of testing and that of attitude research. It found that a new approach to the problem is possible which may help resolve the various conflicting theories and to unify the present opposing schools of thought.

The new approach is that of order factors instead of just common factors. All preceding theories are concerned only with hypotheses about common factors. This turned out to be but a special--indeed, degenerate--case of order factors. When the degeneracy is removed through the concept of a full order, the two basic kinds of factor patterns emerge which we call the simplex and circumplex, respectively. The combination of these two patterns is a full, possible explanation of the workings of the cortex for mental tests and also seems appropriate for many attitude problems. When the simplex-circumplex structure is looked at in its entirety, it is then seen that the present six or seven different approaches to factor analysis have each been examining but one part of the elephant. The major point of each school of thought is found to be correct, although incomplete.

Preliminary analyses of some published tables of intercorrelations of mental tests show that they do approximate the simplex-circumplex theory. The proposed project is to develop more exact computing procedures for the simplex-circumplex approach, and to make more extensive analyses of data already to be found in the literature and already in the files of the Israel Institute.

As part of this project, we shall go over existing materials in order to be better prepared for designing a deeper and more fundamental project of gathering new data in the quantities and design needed fully to establish the simplex-circumplex theory. Dr. Louis Guttman, the Scientific Director of the Israel Institute, who initiated the new approach, has already developed the necessary fundamental mathematics.

APPENDIX II

Personnel Employed Fully or in Part on the Project During the Year

Principal Investigator: Elmo C. Wilson

Project Director: Louis Guttman

Statistician, Research Analysts, and Assistant Statisticians:

Ruth Ludwig
Moshe Sandberg
Judith Tannenbaum
Yaacov Wolf
Mordechai Rosenthal

Clerical Personnel:

Simcha Brudno
Simon Florsheim
Erwin Honig
Reuben Fischer
Rachel Levi
Tova Hari
Laya Beitchman
Rachel Haginguite

APPENDIX III

Breakdown of Actual Costs of the Project

One supervising project director (1/4 time)		$ 1,500.00
One statistician (1/4 time)		1,200.80
One assistant statistician (full time)		4,087.20
One assistant statistician (full time)		3,687.20
Clerical		1,393.60
Supplies, typing, stencilling, etc.		1,098.70
Use charge of three Marchant electric calculators		815.00
New York costs:		
Two research analysts (part time)	$ 965	
Three clerical people (part time)	555	
Miscellaneous	40	
		1,560.00
Total actual costs		$15,342.50
Amount approved by ONR according to original budget estimate		13,315.00
Difference*		$ 2,027.50

* The difference is due largely to the rise in the cost-of-living index.